1962

A Study of Inductive Reasoning Ability Among College Women

Ann Codd Forst
Loyola University Chicago

A STUDY OF INDUCTIVE REASONING ABILITY AMONG COLLEGE WOMEN

by

Ann Codd Forst

A Thesis Submitted to the Faculty of the Graduate School of Loyola University in Partial Fulfillment of the Requirements for the Degree of Master of Arts

June

1962

Life

Ann Codd Forst was born in Evergreen Park, Illinois, on March 15, 1935.

She was graduated from Aquinas Dominican High School, Chicago, Illinois, June, 1952, and from Mundelein College, Chicago, Illinois, June, 1956, with the degree of Bachelor of Arts.

The author began her graduate studies in psychology at Loyola University in June, 1956. At the same time, she became employed as a teacher-technician with the Hearing and Vision Conservation Program at the Chicago Board of Education. The writer retained this position until June, 1958, when she began a clinical internship at Loyola Center for Guidance and Psychological Service where she later became a member of the staff and remained until July, 1960.

Table of Contents

List of Tables

Introduction

In recent times there has been considerable concern about identifying and utilizing scientific aptitudes and interests among our young people. It is hoped that the present study will help to stimulate further interest and research if not in itself contribute modestly to the preliminary investigation of a tool for the identification of an aspect of these talents.

Since it is not the purpose here to engage in philosophical controversy, this writer states her acceptance of the premise that the inductive method is the method which is employed in scientific investigation, and that it is a commonly employed method of making experimental discoveries. If facility in the ability to reason inductively is characteristic of the research scientist, then successful measurement of it may enable students to be guided into areas where their abilities can be used most fruitfully. Investigation of a test of inductive reasoning ability, then, is the major concern of this study. Do scientists differ from others in their inductive reasoning ability? What are some other factors that this ability is related to?

Bittle defines induction as " . . . the legitimate inference of universal laws from individual cases." (1953, p. 342) In a wider sense one is doing inductive reasoning in going from particulars to a pattern or theory, from a group of specific instances to the theory governing them. The derivation of the general "law" or pattern from particular instances is a case of the development of what logicians call "reflex universals."

Epistemologists speak of two types of induction, "complete" and "incomplete." Scientific induction is obviously not the complete induction as defined by epistemologists. Rather it is incomplete because it involves the observation of a limited number of occurrences. As an example, the research scientist sets up an hypothesis. Where there is a sufficient number of instances supporting this hypothesis and no results to the contrary, the hypothesis and the theory become a law.

Spearman, who has contributed so much to the science of psychology, has a classic psychological approach to the process of arriving at knowledge in his formulation of what he calls "Principles of Cognition."

The first cognitive principle, "The Apprehension of Experience," states: "<u>Any lived experience tends to evoke immediately a knowing of its characters and experiencer.</u>" (Spearman, 1927, p. 48) In other words, all that we can know, but infinity in range, depends on the range of the individual person's experience.

In the second principle, "Eduction of Relations," Spearman says, "<u>The mentally presenting of any two or more characters</u> (simple or complex) <u>tends to evoke immediately a knowing of relation between them.</u>" (1927, p. 63) The intellect connects the elements of its experience to attain knowledge.

The third cognitive principle, "Eduction of Correlates," Spearman formulates as follows: "The presenting of any character together with any relation tends to evoke immediately a knowing of the correlative character." (1927, p. 91) The self or intellect, by its very nature, reaches out for new knowledge from that already perceived.

As was mentioned previously, the scientist attains knowledge in the laboratory by means of the inductive method. What this study is concerned with is the reasoning ability that the scientist may possess in his use of the inductive method. In testing for inductive reasoning ability this reasoning factor is found in " . . . tasks that require the subject to discover a rule or principle that covers the material of the test." (Thurstone & Thurstone, 1941, p. 6) Number series, letter series, and other materials are used in testing this ability. The testee is supposed to discover the pattern of the numbers or letters so that he can complete the series.

The instrument used in this study is the Loyola Induction Study (LIS) devised by Charles I. Doyle, S.J. (see Appendix I for sample of LIS). It is a number series test patterned after other tests designed to measure inductive reasoning ability. In this test the subject is required to discover the pattern of the numbers so that he can complete the series with three digits. Having been tested on hundreds of subjects, the items have been scaled for difficulty, and the time limits fixed satisfactorily. This is an investigation preliminary to the standardization of the LIS. Other investigations will be reported in the next section.

The formal hypotheses of this study can now be stated in the form of null hypotheses as follows:

1. There is no difference in inductive reasoning ability between upper-division science majors and humanities majors in a women's college as judged by performance on the LIS.

2. There is no relationship between upper-division college women's performances on a test of inductive reasoning, the LIS, and scholastic ability as measured by a college entrance test.

3. There is no relationship between upper-division college women's performances on a test of inductive reasoning, the LIS, and scholastic achievement as measured by college grade point average.

Review of Related Literature

Since the material of this study is a number series test of inductive reasoning ability, this review will attempt to cover a representative selection of studies on the development of such tests, their discriminative value among academic interests, and their relationships to scholastic ability and achievement. In some instances the inductive reasoning or number series test forms but part of a larger division of a test sampling varied skills or abilities.

The Army Alpha Test was developed for testing recruits during the First World War. It contained a number series subtest developed from Miss Rogers' missing number test, and was one of the first times that this method was used in testing mental ability. F. L. Wells collaborated with others in producing the number series subtest. The first revision of the examiner's guide for the test reported that the coefficient of correlation of seven other Alpha subtests with the number series subtest ranged from .61 to .77, and the correlation with total score was .84.[1] (Yerkes, 1921)

Whipple (1927) used the sixth form of the Army Alpha with 227 boys and 317 girls in their freshman year and 129 boys and 161 girls in their senior year of high school. In the number series subtest the freshman and senior boys' median score points exceeded the girls' points by 0.79 and 0.88 respectively. The author notes this sex difference but remarks upon the sexless content of the number series subtest.

In studying verbal and numerical abilities, Schneck (1927) used 210 college men of Hebrew extraction ranging in age from 18 to 21. Nine tests were administered. One of the numerical subtests was a number series test which was a forty-item power test requiring a two-number answer. Four of the items were from the Army Alpha and the rest were the author's own. The highest coefficient of correlation for the number series subtest was .92 with the arithmetic reasoning subtest. With age partialled out and after correction for attenuation, it still correlated highest with arithmetic reasoning at .46. The coefficient of correlation between the number series subtest and the numerical division of the battery was .46.

L. L. Thurstone (1919) included a number series subtest in his Psychological Examination for College Freshmen and High School Seniors in 1919.

Thurstone asserted that particular activities involve various combinations of abilities. He utilized the method of factor analysis to arrive at what he called his "primary mental abilities." The number of these primary abilities has varied from study to study.

In the early 1930's, Thurstone (1946) reported the isolation of twelve factors from a battery of 56 tests administered to several hundred college students. An inductive reasoning factor was among these. He felt that engineering and science students were characterized by a combination of visualizing and reasoning factors.

Thurstone (& Thurstone, 1941) acknowledged difficulty in obtaining high validities in testing the inductive reasoning factor, but felt quite confident that it did exist. The method of testing was to have the subject find the rule governing the test material. This factor has been found in tests of many different types, so induction seems to "transcend" the material of the test.

In order to see if primary abilities could be isolated for the fourteen-year-old level, Thurstone (& Thurstone, 1941) reported that a battery of sixty tests was administered to 1154 children in grade 8-B. The completed records of 710 children were available and so were used for this study. Eight variables had principal loadings on the induction factor. From the results of the study, 21 tests were selected to be given to 437 children in eighth grade and the upper half of seventh grade in December, 1939. The primary factors which seemed to be apparent enough to be considered in this study were: V (verbal comprehension), W (word fluency), N (number), S (space), M (rote memory), I preferably called R (induction or reasoning), and P (perceptual). The P factor was not too definite however. The D (deduction) factor was not used in this study. The three induction tests chosen for this study included Letter Series, Pedigrees, and Letter Grouping, because they had the highest loadings for I in the previous study. In this study they had loadings on the R factor of .53, .44, and .42 respectively. Thurstone felt that these loadings were not as high as might be desired, but the tests were retained because the R factor correlated .84 with the second order general factor which was the highest loading on this factor in the correlated primaries. He felt that this might be close to what Spearman called "g."

The experimental test edition of Thurstone's Primary Mental Abilities tests (PMA) came out in 1938, with the measurement of seven primary abilities. Then the Chicago Tests of Primary Mental Abilities were published in 1941, by the American Council on Education. Six primary abilities were tested in this edition. At present there are batteries available for different age levels.

Ellison and Edgerton (1941) used the experimental edition of the PMA with 49 students at Ohio State University to investigate the usefulness of the primary abilities for scholastic guidance. The D factor in this edition contained a number series subtest, while the I factor contained a letter grouping subtest. The number series subtest correlated .35 with grade point average and .36 with intelligence as judged by the Ohio State University Psychological Examination.

Goodman (1943) also used the experimental edition of the PMA. With 170 freshman engineering students at Penn State College he found a Pearson product-moment coefficient of correlation of .38 between the letter grouping subtest of the I factor and the number series subtest of the D factor.

Hobson (1947) studied sex differences in the PMA. Using the 1941 edition with 222 boys and 250 girls in ninth grade, he found a CR of 5.45 for the means on the R subtest favoring the girls. He then used the revised single-booklet edition with 265 boys and 260 girls in ninth grade and found a CR of 4.53 for the R subtest favoring the girls. Using 720 eighth grade boys and 716 girls, a Cr of 7.45 for the R subtest favoring the girls was found. For these same groups the CR for IQ difference on the Kuhlmann-Anderson Test favored the girls by yielding 1.78, 3.67, and 4.49 respectively. The author said that differences in the specific factors couldn't be completely attributed to the higher intelligence shown by the girls, since there was a greater difference for the PMA R subtest than for the Kuhlmann-Anderson. He attributed the higher mean IQ to the fact that the Kuhlmann-Anderson has a heavy loading on N, W, and R in which the girls were found superior. The author felt that the PMA R subtest was restricted and of a type which would be easier and more appealing to girls. He concluded that, for the PMA to be used for guidance, separate age and sex norms for the subtests would be needed.

Botzum (1951) reported his analysis of 46 tests given to 237 college men in order to study the Reasoning and Closure factors. He remarked that Thurstone's I factor has proved to be the most constant of the reasoning factors in other studies. Botzum used the product-moment method of correlation and the split-half reliability method. Thurstone's complete centroid method was used in factoring the correlation matrix twice. Among others, four of the highest loadings on the I factor included letter series .47, number series .45, letter grouping .38, and number patterns .36. The reliability of the number series subtest was .77. Its highest correlation was .65 with the letter series subtest. On the first second order factor the author found loadings of .74 for S, .68 for D, .67 for I, and .64 for flexibility of closure. Thurstone had found somewhat similar results. Botzum felt that one way to explain the results was that either the synthetic or analytic methods could be used to solve closure and space problems. He felt another might be that both D and I have certain gestalt elements.

Segel (1944) used the Office of Education Scientific Aptitude Test, a prognostic test in math and science for the upper high school level. It contained a number series subtest. Using the split-half method a reliability of .87 was obtained with the battery for sixty boys in math and science classes in senior high school. Correlations between the battery and subject marks averaged .41 with science for 423 students, .40 with math for 355 students, and .29 with English for 83 students. Segel concluded that the battery as a prognostic instrument did discriminate between academic subjects and math and science courses.

In 1946, Baldwin (1946) developed his Inductive Reasoning Test, a number series test for the ninth through twelfth grades. Coefficients of correlation between test scores and teachers' rankings averaged .62. There was a correlation of .69 between the scores and science and math senior high school grade points.

The Kuhlmann-Anderson Tests have an induction subtest. The manual (Personnel Press, Inc., 1952) reported that, for one hundred ninth grade pupils, the induction subtest correlated .74 with the total of the other subtests.

Thus far the literature and studies reviewed have been concerned with number series tests and subtests. Now the reader's attention is directed toward the criterion of scholastic ability that was used in this study. This criterion was the American Council on Education Psychological Examination for College Freshmen (ACE). It is considered to be a highly reliable test.

The American Council on Education introduced into the 1938 edition of the general Psychological Examination two divisions. One division was made up of three Linguistic subtests, and the other contained three Quantitative subtests. These subtests were not supposed to represent the primary mental abilities, but two groups of abilities as their names designated. They were advanced as a "compromise" with the PMA tests by introducing the binary division and a shorter testing time. (Thurstone & Thurstone, 1947)

Thurstone was primarily responsible for the earlier editions of the examination for the American Council on Education. The Cooperative Test Service of the American Council became the Cooperative Test Division of the Educational Testing Service in 1948, and thus became responsible for publishing the test. In the ACE, the Linguistic (L) subtests include: Same-Opposite, Completion, and Verbal Analogies. The Quantitative (Q) subtests include: Arithmetic, Number Series, and Figure Analogies. (Educational Testing Service, Cooperative Test Division, 1953)

In one study on the ACE (Anderson, Anderson, Ferguson, Gray, Hittinger, McKinstry, Motter, & Vick, 1942) the 1940 form of the test was administered to 112 freshman college women. Six weeks later they were given the 1941 form. The grade point averages for the year correlated .48 and .36 respectively for the L and Q scores of the 1940 edition. They correlated .54 and .39 respectively for the L and Q scores of the 1941 form. The correlation between L and Q in the 1940 form was .61 and, on the 1941 form, it was .45. They also found that the L score had as high a correlation with grade point average as the Total score. As a result, the authors stated their doubt as to the value of the Q score in predicting scholastic achievement and felt that a briefer version of the ACE might be just as effective as the present form appeared to be.

MacPhail (1942) studied the predictive ability of the Q and

L scores of the ACE. For 52 engineering majors, 25 chemistry majors, and 324 liberal arts majors, he found that the L score, with a CR of 4.88, was a better predictor of freshman year averages for the liberal arts group. There had been no significant difference for first semester averages. Then he considered the first year average grades for quantitative studies, including psychology, using 349 students and for verbal studies using 398 students. There was no significant difference between Q and L for quantitative studies. But L was significantly better than Q in predicting the first year average for verbal studies with a CR of 4.80. Then individual courses were correlated with the Q and L scores. In general, neither score proved much more effective than the other. As a result, the author concluded that counselors should be very careful in using the Q and L scores of the ACE in differential prediction.

Thomann (1948) reported a study using 566 boys and girls in correlating first year grade point average at the University of Illinois with the 1940 college edition of the ACE. Correlations with grade point averages were .413 for the L score, .316 for the Q score, and .443 for the Total score of the ACE.

Barrett (1952) reported a study of ACE scores for differential prediction in college math. With the 1943 edition, correlation differences of .21 and .185 between the Q and L scores favored the Q scores in the prediction of freshman grades in trigonometry and college algebra respectively. A correlation difference of .202 between the Q and Total ACE scores favored the Q score in predicting trigonometry grades. These differences were significant. Other differences with the 1943 edition or the 1947 or 1948 editions were not significant or consistent. Barrett felt that the Q score was not consistently better than L in predicting math achievement, and consequently, shouldn't be used in that way.

The review now brings us up to the instrument which is the focal point of this study, the LIS. After a brief review of its development, some of the preliminary investigations of the LIS will be examined.

In a laboratory experiment, C. I. Doyle, S.J. (1933) investigated inductive discovery by means of an original, multiple-choice, self-recording keyboard. The problem for 31 subjects was to induce a bell to ring in as few strokes of the keys as possible. False hypotheses were often formed and held to because of the interposition of a time switch. However, a significant improvement was observed with subjects who were instructed to apply the scientific method to the problem, ie., alert observation, accurate notes, and making a distinction between the hypothesis and observed facts. The results of the study indicated that the trials were governed by thought, thus setting inductive discovery apart from trial-and-error learning.

Later the same experimenter devised his paper-and-pencil test of inductive reasoning ability known as the LIS. It is a number series test as described in the Introduction.

McNeill (1959) used the LIS with one group of one hundred bright male high school students in their junior and senior years, the majority of whom were interested in science. The other group consisted of eighty male and twenty female graduate and undergraduate university students who were primarily interested in industrial relations. The high school mean on the LIS was 45.62 with a SD of 9.22. The adult mean was 37.62 with a SD of 12.57. The difference between the means was significant at the .01 level of confidence. The experimenter concluded that the higher high school mean may have been due to the fact that it was an accelerated group of superior intelligence, since inductive ability is closely related to intelligence.

Sister Mary Colomana Buksa (1960) used 212 girls and 188 boys from grades five through eight as subjects for 42 items of the LIS. The fifth grade mean was 13.34 with a SD of 7.44. The sixth grade mean was 16.70 with a SD of 5.91. The mean for seventh grade was 18.98 with a SD of 8.04. In eighth grade the mean was 23.30 with a SD of 7.17. The differences between the fifth and sixth grades and between the seventh and eighth grades in mean scores were both significant at the .001 level of confidence. The difference for the sixth and seventh grades was significant at the .05 level. Some accident of sampling yielded median scores which were more significant than the mean scores. These scores for grades five through eight were 13, 16, 19, and 22 respectively. The author concluded that the ability for inductive reasoning develops with age and can be measured reliably.

In a study somewhat parallel to the one undertaken in this paper, V. Ortolani, S.J. (1959) reported the use of 128 male college upper-classmen evenly divided between science and humanities as their major fields in order to see if inductive reasoning ability is related to scientific ability. The mean for the science group on the LIS was 44.5 with a SD of 10.7. The humanities' mean was 39.8 with a SD of 10.6. The difference between the means was significant at the .02 level of confidence. However, on admission to college, a higher percentile cutting point was required for those electing a science curriculum. But, when a smaller group of subjects was matched for ACE Total score and the means recalculated, the results were not significant. The mean grade point average for the humanities group was 2.39 with a SD of .435. The science mean was 2.63 with a SD of .438. The difference was significant at the .01 level of confidence. When 32 from each group were matched for grade point average and the means recalculated, the difference favoring the science group was significant at the .02 level of confidence. The experimenter then determined the influence of quantitative ability from the ACE upon LIS scores by computing the Q/L ratio for available subjects and then correlating the ratio with the LIS. Rank-difference coefficients of correlation of .41 and .49 were obtained for 44 science and 35 humanities majors. The product-moment coefficient of correlation for the entire group was .42. The author concluded that inductive reasoning ability is characteristic of, but not peculiar to, science majors in general. It was also stated that inductive reasoning ability does not seem to influence the attainment of humanities students in their own subjects or in the subjects which are common to both groups.

Procedure

The subjects taking part in this experiment consisted of junior and senior students in a college for women. They were divided into two groups of 64 students each. The group designated as "science" consisted of 42 juniors and 22 seniors with majors in the following departments: 21 from psychology, 19 from biology, 15 from chemistry, and 9 from mathematics. Their mean age was 20.8.

The group designated as "humanities" consisted of 45 juniors and 19 seniors with their major fields as follows: 21 from English, 13 from history, 9 from sociology, 6 from speech, 4 from art, 4 from economics, 3 from drama, 3 from home economics, and 1 from French. Their mean age was 20.4.

The subjects were tested in class groups and in large and small volunteer groups. When the students were given the LIS test booklet, they were asked to fill out the personal data at the top of the first page and then to stop. The directions were uniform for all administrations of the test. The subjects were reminded that this is not a test of intelligence, but a study of how people make discoveries. They were then requested to complete with three digits each of the four sample series of numbers and then to stop. When everyone was finished and understood the procedure, they were instructed to turn the page, and the timing began. At the end of twenty minutes the tests were collected. The raw score was the number of series which were entirely correct including the sample series. Although this wouldn't have happened if the students had followed the directions, six subjects did not complete all of the four samples, including one who also began the answer to a series correctly but did not complete it. These sample series were counted as correct since later more difficult series were done correctly. There were also eight subjects who had errors among the four sample series. They were not given credit for these.

The criterion of scholastic ability in this study was the ACE which was used as part of the entrance testing at this particular college. Raw scores for the L and Q subtests and the Total test were available for 105 subjects, 49 science majors and 56 humanities majors. Seventy students had taken the 1952 form of the test, and 35 had taken the 1946 form. Equivalent scores for the two forms were not available. However, the effects of the 1946 form may be somewhat balanced since, of the 35 subjects who took it, 18 were from the science group and 17 from the humanities group.

The criterion that was used for scholastic achievement was the grade point average. Only a student's grades obtained at the present college were used, and averages for at least four full semesters were required. This college point system credits three points per semester hour for an "A," two points for a "B," one point for a "C," and none for a "D." Grade point averages were rounded off to the nearest hundredth. These averages were available for 105 students, 49 from the science group and 56 from the humanities group. Forty-eight of the science majors and 54 of the humanities majors used for this part of the study were the same subjects included in the ACE study.

In brief, the statistical procedures that were employed in this study were as follows:

1. The mean LIS scores for the science and humanities groups were determined, and then the significance of the difference between the means of the two groups was computed.

2. LIS scores were correlated with available Q, L, and Total scores of the ACE for the science and humanities groups and both groups combined.

3. LIS scores were correlated with the Q/L ratio of the ACE for the science, humanities, and combined groups.

4. LIS scores were correlated with available grade point averages for the science, humanities, and combined groups.

Results and Discussion

The means, standard deviations, the difference between the means, and the critical ratio were obtained for the LIS scores of a group of 64 upper-division science majors and a group of 64 upper-division humanities majors in a women's college. The results are shown in Table 1.[2]

Table 1

Comparison of Mean LIS Scores for Two Groups of Upper-Division College Women

Group	Number	Range of Scores	Mean	SD
Science	64	20-61	43.59	10.41
Humanities	64	12-53	35.73	8.35

Difference between Means = 7.86   CR = 4.71

P = .001

The difference between the means for the two groups on the LIS was significant at the .001 level of confidence, favoring the science group. So the first null hypothesis of this study may be rejected with a high level of confidence, i.e., there is a difference in inductive reasoning ability between upper-division science and humanities majors in a women's college as judged by performance on the LIS. Since this was a random sampling of subjects, it can be said that, in general, the science majors are superior in inductive reasoning ability to the humanities majors judging from performance on the LIS.

The next step in this study involved determining the relationship between inductive reasoning ability and scholastic ability. The LIS scores for the science, humanities, and combined groups of upper-division college women were correlated with the Q and L subtests and Total score of the ACE. The product-moment method was used in obtaining the coefficients of correlation for this comparison and all of the following in this section.[3] To assess sampling errors for all correlations in this section the standard error of r was used.[4] The results appear in Table 2, page 23.

All of the correlations between the ACE and the LIS were positive. So the results shown in Table 2 reject the second null hypothesis since they reveal a relationship between upper-division college women's performances on a test of inductive reasoning, the LIS, and scholastic ability as measured by a college entrance test.

The combined group and the science group, in that order, had higher correlations between the LIS and the ACE Total score than did the humanities group. There were very little differences between the groups in the correlation between LIS scores and the ACE L score. The highest correlation reported in Table 2 was that which was found between the LIS and the ACE Q score for the science group.[5] It was considerably higher than that obtained for the humanities group, but it approximated the correlation obtained for the combined group between the LIS and the Q score.

Table 2

Coefficients of Correlation Between LIS Scores and Q, L, and Total Scores of ACE for 105 Upper-Division College Women

ACE	LIS					
	Science (N 49)		Humanities (N 56)		Combined Group (N 105)	
	r	σ_r	r	σ_r	r	σ_r
Q score	.726	.144	.571	.135	.723	.098
L score	.378	.144	.394	.135	.409	.098
Total score	.592	.144	.545	.135	.626	.098

Over and above the rejection of the null hypothesis, it can be said that there is definitely a higher relationship between the Q subtests of the ACE and inductive reasoning ability as shown on the LIS than either the L or Total score relationships. The higher relationship between the LIS and ACE Q score in the case of the science group, and consequently the combined group, might possibly be explained in four different ways.

The first explanation for the high relationship could be that both the Q score of the ACE and the LIS have some factor in common. This could be that the LIS is a number series test and one of the subtests of Q on the ACE is a number series test. Intelligence might also be a common factor. It is not within the scope of this study to investigate relationships of this type any further.

The second explanation for the high relationship might be that good inductive reasoning ability contributed to good performance on the quantitative subtests. As shown in Table 1, the science majors, in general, do appear to be superior in inductive reasoning ability to the humanities majors as judged by their performance on the LIS.

A third explanation for the high relationship between the LIS and the Q score of the ACE for the science, and consequently the combined groups, might be that high quantitative ability contributed to proficiency in inductive reasoning. In order to test the influence that Q might have had on LIS performance, the Q score was divided by the L score and this ratio determined for all of the subjects who had taken the ACE. Then this Q/L ratio was correlated with the respective LIS scores for the science, humanities, and combined groups. The results of these correlations are presented in Table 3, page 25.

All of the coefficients of correlation shown in Table 3 are positive. But the results were contrary to expectations since the comparisons of the LIS scores with the Q/L ratio of the ACE did not reveal a high correlation for the science group. Thus Q, or quantitative ability, did not appear to have had the expected influence on inductive reasoning for the science group. The ACE Q score had greater influence on LIS performance for the humanities group apparently, and thus, an even greater influence on the combined group in LIS performance. So the third explanation, that high quantitative ability contributed to proficiency in inductive reasoning, does not appear to be plausible, at least insofar as these results indicate.

Table 3

Coefficients of Correlation Between LIS Scores and Q/L Ratio of ACE for 105 Upper-Division College Women

ACE	LIS					
	Science (N 49)		Humanities (N 56)		Combined Group (N 105)	
	r	σ_r	r	σ_r	r	σ_r
Q/L ratio	.330	.144	.384	.135	.480	.098

A fourth explanation for the high relationship between the LIS and the ACE Q score for the science group could involve the motivation of the two academic groups. It is a possibility that science majors, as a group, might have more definite career plans than the humanities majors, and therefore, are more serious in their academic attitudes. The humanities majors might not appear to be as selective a group as the science students because their career plans might not be as definite. Some may merely be filling in time between high school and marriage, perhaps desiring a degree in order to obtain a more appealing job before marriage. This explanation is, of course, merely hypothetical, since the writer has no objective data about the subjects to substantiate it.

The last step in these analyses concerns the investigation of the relationship between the LIS and scholastic achievement. The LIS scores for the science, humanities, and combined groups of upper-division college women were correlated with available grade point averages for at least four full semesters. The results appear in Table 4.

Table 4

Coefficients of Correlation Between LIS Scores and Grade Point Averages for 105 Upper-Division College Women

Scholastic Ability	LIS					
	Science (N 49)		Humanities (N 56)		Combined Group (N 105)	
	r	σ_r	r	σ_r	r	σ_r
Grade point average	.317	.144	.473	.135	.491	.098

The coefficients of correlation obtained in Table 4 were all positive. The third null hypothesis of this study may then be rejected since the results reveal a relationship between upper-division college women's performances on a test of inductive reasoning, the LIS, and scholastic achievement as measured by college grade point average.

The humanities and combined group were close in their relationships between the LIS and grade point average. The science group showed a less evident relationship, suggesting that scholastic achievement is less dependent upon inductive reasoning ability for them than it is for the humanities group. This low correlation was unexpected. Any attempt to explain it would be pure speculation.

Comparing parts of this study with Ortolani's study (1959), the parallel group of women subjects used here showed a significant mean difference on the LIS favoring the science group at a higher level of confidence than the men did. Although using a smaller group of men in comparing LIS scores with the Q/L ratio of the ACE, Ortolani's results were lowest for the science majors as this study also showed, and thus did not bear out the experimenters' expectations.

Summary and Conclusions

This study began by accepting the premise that the inductive method is used in scientific discovery, but raised the question of whether scientists differ from others in the ability to reason inductively in applying this method. Bittle's definition of induction as " . . . the legitimate inference of universal laws from individual cases," was given. Inductive reasoning, according to Thurstone, is found in " . . . tasks that require the subject to discover a rule or principle that covers the material of the test." The test used in this study, the Loyola Induction Study (LIS), is composed of 62 number series items requiring three digits in response. The null hypotheses for the study were then stated as follows:

1. There is no difference in inductive reasoning ability between upper-division science majors and humanities majors in a women's college as judged by performance on the LIS.

2. There is no relationship between upper-division college women's performances on a test of inductive reasoning, the LIS, and scholastic ability as measured by a college entrance test.

3. There is no relationship between upper-division college women's performances on a test of inductive reasoning, the LIS, and scholastic achievement as measured by college grade point

average.

Some studies were reviewed concerning the relationship of number series tests or subtests to academic interests, abilities, and achievements.

The experimental procedure involved the administration of the LIS to a group of 64 upper-division women science majors and to a group of 64 humanities majors. LIS raw scores were used. The American Council on Education Psychological Examination for College Freshmen (ACE) was used as the criterion of scholastic ability. Grade point averages were used as criteria for scholastic achievement.

Table 1 showed a comparison of the mean scores of the LIS for both groups of 64 subjects each. The science group was significantly higher at the .001 level of confidence, and thus, the first null hypothesis was rejected.

Table 2 showed the correlation between LIS scores for the two groups separately and combined with available Quantitative (Q), Linguistic (L), and Total scores of the ACE. The second null hypothesis was rejected since all of the correlations were positive. The highest correlation reported in Table 2 was .726 between the Q score of the ACE and the LIS score for the science group.

One explanation for this high relationship was that both the ACE Q score and the LIS may have a common factor. The second explanation was that high inductive ability (LIS) contributed to high quantitative (Q score) performance. This viewpoint was strengthened by the higher mean LIS score for the science group as indicated in Table 1. The third explanation advanced was that high quantitative performance (Q score) contributed to high inductive reasoning performance (LIS).

Table 3 showed the Q/L ratio (Q score divided by L score) of the ACE correlated with LIS scores to assess the influence of quantitative ability on inductive reasoning ability. The science group's coefficient of correlation of .330 was the lowest shown. Therefore, quantitative ability (Q score) did not appear to have had the expected influence upon inductive reasoning performance (LIS) for this group.

A fourth explanation advanced for the high correlation between the Q score of the ACE and the LIS score for the science group was that the science group might be more highly motivated because of more definite career plans than the humanities group.

Table 4 showed correlations of grade point averages with the LIS for the two groups separately and combined. All of the resultant correlations were positive, so the third null hypothesis was rejected. The science group showed the lowest correlation between grade point average and the LIS at .317, with the correlation for the humanities group being .473. No explanation for this could be given.

By comparison, the mean LIS differences favoring the science women were in the same direction but more significant than those of Ortolani's (1959) parallel group of men. His results with the Q/L ratio of the ACE were in the same direction as this study.

From the foregoing study, the following conclusions result:

1. In general, upper-division science majors appear to be superior to humanities majors from this women's college in inductive reasoning ability as judged by performance on the LIS.

2. There is a positive relationship between upper-division women's performances from this college on a test of inductive reasoning, the LIS, and scholastic ability as measured by a college entrance test, the ACE.

3. These upper-division women science majors show a high relationship between quantitative ability, judged by the Q score of the ACE, and inductive reasoning ability, as judged by performance on the LIS. This may be due to higher inductive reasoning ability on the part of the science students, or to a factor common to both tests, or to more serious academic interests among the science majors. The relationship does not appear to be due to the influence of quantitative ability on inductive reasoning.

4. There is a positive relationship between upper-division women's performances from this college on a test of inductive reasoning, the LIS, and scholastic achievement as measured by college grade point average.

5. No immediate explanation can be offered for the higher relationship between LIS scores and grade point averages for the women humanities majors as compared with the science majors.

References

Anderson, E. E., Anderson, S. F., Ferguson, C., Gray, J., Hittinger, J., McKinstry, E., Motter, M. E., & Vick, G. Wilson College studies in psychology: I. A comparison of the Wechsler-Bellevue, Revised Stanford-Binet, and American Council of Education tests at the college level. _J. Psychol._, 1942, 14, 317-26.

Baldwin, G. B. _Inductive Reasoning Test._ Minneapolis: Educational Test Bureau, 1946.

Barrett, Dorothy M. Differential value of Q and L scores on the ACE Psychological Examination for predicting achievement in college mathematics. _J. Psychol._, 1952, 33, 205-07.

Bittle, C. N., O.F.M. _Reality and the mind: epistemology._ Milwaukee: Bruce, 1953.

Botzum, W. A. A factoral study of the reasoning and closure factors. _Psychometrika_, 1951, 16, 361-86.

Buksa, Sister Mary Colomana, C.S.S.F. Preliminary study of a test of inductive reasoning in the upper grades. Unpublished master's thesis, Loyola Univ., Chicago, 1960.

Doyle, C. I., S.J. An experimental investigation of the process of inductive discovery with groups of closely similar problems of variable complexity. Unpublished doctoral dissertation, St. Louis, 1933.

Educational Testing Service, Cooperative Test Division. Norms bulletin: American Council on Education Psychological Examination for College Freshmen, 1952 edition. Princeton, N. J.: Author, 1953.

Ellison, M. L., & Edgerton, H. A. The Thurstone Primary Mental Abilities Tests and college marks. Educ. psychol. Measmt., 1941, 1, 399-406.

Goodman, C. H. A factorial analysis of Thurstone's sixteen primary mental abilities tests. Psychometrika, 1943, 8, 141-51.

Hobson, J. R. Sex differences in primary mental abilities. J. educ. Res., 1947, 41, 126-32.

McNeill, Mary. A study of the relation between deductive and inductive reasoning ability and adjustment in adults. Unpublished doctoral dissertation, Loyola Univ., Chicago, 1959.

McNemar, Q. Psychological statistics. New York: John Wiley & Sons, 1949.

MacPhail, A. H. Q and L scores on the ACE Psychological Examination. Sch. & Soc., 1942, 56, 248-51.

Ortolani, V., S.J. A preliminary study of differences in inductive reasoning ability among upper division college men. Unpublished master's thesis, Loyola Univ., Chicago, 1959.

Otis, A. S. Otis correlation chart. Yonkers-on-Hudson: World Book, 1922.

Personnel Press, Inc. _Master manual_: _Kuhlmann-Anderson Tests._ (6th ed.) Princeton, N. J.: Author, 1952.

Schneck, M. M. R. The measurement of verbal and numerical abilities. _Arch. of Psychol._, 1929, 17, No. 107.

Segel, D. Validity of the V. C. Aeronautics Aptitude Test and the O. E. Scientific Aptitude Test. _J. Psychol._, 1944, 18, 65-80.

Spearman, C. _The nature of 'intelligence' and the principles of cognition._ London: Macmillan, 1927.

Thomann, D. F. Relationships between the high school and college editions of the American Council on Education Psychological Examination and their relative value in predicting college achievement. _Coll. & Univ._, 1948, 23, 217-33.

Thurstone, L. L. _Psychological Examination for College Freshmen and High School Seniors._ (Copy). Pittsburgh: Carnegie Instit. of Tech., 1919.

Thurstone, L. L. Theories of intelligence. _Sci. Mon._, _N. Y._, 1946, 62, 101-12.

Thurstone, L. L., & Thurstone, T. G. Factorial studies of intelligence. _Psychometric Monogr._, 1941, No. 2.

Thurstone, L. L., & Thurstone, T. G. _Manual of instructions: American Council on Education Psychological Examination for College Freshmen, 1947 edition._ Washington: American Council on Education, 1947.

Whipple, G. M. Sex differences in Army Alpha scores in the secondary school. _J. Ed. Res._, 1927, 15, 269-75.

Yerkes, R. M. (Ed.) _Memoirs of the National Academy of Sciences._ Vol. XV. _Psychological examining in the United States Army._ Washington: Government Printing Office, 1921.

Footnotes

[1] All of the correlations cited in the Review of Related Literature are positive.

[2] For statistical procedures used in Table 1, see McNemar (1949, pp. 16, 25, 65, 223, and 352).

[3] The product-moment method of obtaining the correlation coefficients was used by means of the Otis Correlation Chart (Otis, 1922).

[4] See McNemar (1949, p. 122) for the formula for standard error of r.

[5] A further study on the Q score of the ACE and LIS performance is presented in Appendix II.

Appendix I

Loyola Induction Study Test

LOYOLA INDUCTION STUDY

Name__________________________ Date__________

Student at____________________________________

Highest year of school completed (circle one)

6 7 8 9 10 11 12 13 14 15 16 _____

What is your favorite study or your major field?____________________________________

INSTRUCTIONS

This is not an intelligence test. It is part of a study of how people make discoveries.

There are some easy examples below. Please read each row of figures and then write in the three blank spaces at the end of each row the numbers that should follow.

2	4	6	8	10	12	___	___	___
9	8	7	6	5	4	___	___	___
1	7	2	7	3	7	___	___	___
2	2	3	3	4	4	___	___	___

N.B. Please do not turn this page until you are told to do so.

Appendix II

Comparison of ACE Q Scores Within Two Groups

A further study of ACE Q scores and LIS performance was made by first, listing the upper-division college women of the science and humanities groups in descending order based on LIS scores; second, determining the means of the available Q scores for the upper and lower portions of each group; and, third, obtaining the significance of the difference between the mean Q scores of the upper and lower portions of each group. The results appear in Table 5. See McNemar (1949, pp. 16, 25, 65, 223, and 352) for the statistical procedures used in Table 5.

Table 5

Comparison of Mean ACE Q Scores for a Binary Division of Two Groups of College Women Based Upon LIS Scores

Group	Upper Portion			Lower Portion			D_M	CR
	N	Q Score Mean	SD	N	Q Score Mean	SD		
Science	24	54.17	5.74	25	43.46	8.22	10.71	5.30*
Humanities	28	44.46	8.75	28	33.79	9.42	10.67	4.39*

* P = .001

The results in Table 5 indicate that the upper-halves of both groups of upper-division college women have significantly higher mean Q scores at the .001 level of confidence than the lower halves of the groups. In the case of the science group, these results confirm the coefficient of correlation of .726 found between the LIS scores and the Q score of the ACE. (Table 2)

Although the humanities group also showed significant differences within its own group, the mean Q scores were lower than those of the science. This seems to indicate the same trend as that shown with LIS mean scores where the science group had a significantly higher LIS mean than the humanities. (Table 1)

APPROVAL SHEET

The thesis submitted by Ann Codd Forst has been read and approved by three members of the Department of Psychology.

The final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated, and that the thesis is now given final approval with reference to content, form, and mechanical accuracy.

The thesis is therefore accepted in partial fulfillment of the requirements for the Degree of Master of Arts.

June 6, 1962
Date

Charles P. Doyle, S.J.
Signature of Adviser